

An application of SGA: How urban resident acts in making traveling decisions

Long Chenxu¹, Li Jing^{1,2}, Li Zhuyi³

¹Beijing Jiaotong University, ²University of Illinois (China), ³Urbana Champaign (USA)

longlonglonglonger@163.com; jingli@bjtu.edu.cn; zhuyili@yahoo.com

Received: August 2012

Accepted: February 2013

Abstract:

Purpose: To study how urban resident acts in making traveling decisions.

Design/methodology/approach: Simple Genetic Algorithm (SGA)

Findings: Using SGA to make the model of urban resident's decisions is rational.

Research limitations/implications: This study is just about single one and the SGA could be made better.

Originality/value: Use SGA to describe urban resident's traveling behavior, and combine the management problem with the AI.

Keywords: urban resident, traveling decision, SGA

1. Introduction

Professor J.H. Holland and his colleagues from University of Michigan, Ann Arbor came up with Genetic Algorithm in the 1860s and then developed it. By the early 1870s, Holland put forward the schema theorem which was regarded as the fundamental theorem of the Genetic Algorithm, consequently, it established the theoretical foundation of the research of Genetic Algorithms.

In the ordinary course of events, Genetic Algorithm convert the unsolved optimization problem to the maximization problem of a fitness function. However, when the display of the fitness function is difficult to show, the Interactive Genetic Algorithm (IGA) emerge at the right moment (Song, Zhu & Wu, 2009). Interactive Genetic Algorithm is a expansion of the traditional Genetic Algorithm, and there are many studies about IGA. Interactive Genetic Algorithm is relatively complex, because of the selecting after every generation, it is very easy to cause the tiredness of users. As a result, there appeared some researches about machines replacing users (Yao, Guan & Sun, 2006).

This paper put forward a modeling of the description of urban residents' traveling decision based on Simple Genetic Algorithm (SGA), and checked the application result with the use of Mat-lab and toolbox from the University of Sheffield.

With the development of economy and urbanization, urban traffic in our country face the greater pressure, urban residents can choose various means of transportation to go out. With the technology of information went deep into the traffic, urban residents would take convenience, economy, safety and comfort into deep consideration before making the trip decisions. How to present the journey wish of the urban residents had been researched (Xu, Li, Zhu & Shi, 2005).

If the urban resident has a clear judgment on the alternative trip modes when he/she decides which transportation to take, it is called urban resident making decision in the determined environment. Conversely, it is called urban resident making decision in the uncertain environment. Ma, Wen, Li and Li (2009) used the double base method to study how urban residents making decisions under determined environment and analyzed how urban residents making decisions under uncertain environment with interval number ranking method.

There are many researches about how urban residents behave when choosing trip modes, it cannot be denied that many of them offer important reference to this paper, especially the factor analysis about how urban residents choose transportation and the travel structure. However, there are few studies discussing the application of the urban residents making trip modes decisions based on Genetic Algorithm, and this is where the innovation of this paper locates.

One of the most important points of this paper is how to describe the urban residents making trip modes decisions based on simple Genetic Algorithm. What's more, to establish a proper fitness function is another difficulty of this paper.

However, the study of individual residents' choice behaviour of making trip modes decision is not enough. We should start with a single individual, and do more research on a group of individuals whom can be divided by locations or incomes. This paper may provide theoretical basis for future study, such as how to select modes of transportation for groups of residents. And then offer theoretical basis for government planning urban transport system, optimizing urban traffic management and solving the traffic congestion. Besides, this paper may provide theoretical support for the network platform built between government and residents. All these researches may lead to a more scientific management on urban traffic system. And this is exactly the meaning and value of this paper.

2. The basic process of Genetic Algorithm

The genetic manipulation of genetic algorithm in the whole evolutionary process is random, but the characteristics it shows are not entirely random search, it can effectively use historical information to speculate that the next generation expect improved performance to find the advantages set. Evolving from generation to generation, finally it converges to a best adapted individual to the environment, and seeks the optimal solution of the problem. The genetic algorithm involves five elements: parameter encoding, setting of the initial population, the design of the fitness function and genetic manipulation, setting proper controllable parameter.

The usually used workflow and structure of the genetic algorithm is first proposed by Goldberg in the natural gas pipeline control optimization applications, commonly known as the regulated genetic algorithm (the canonical GA, CGA) or the standard genetic algorithm (standard GA, SGA).

The basic processes and structures of the simple genetic algorithm are shown in Figure 1 and Figure 2.

From Figure 1 and Figure 2, it shows that the genetic algorithm is running as a typical iterative process, the work must be completed and the basic steps are as follows:

(1) Select the encoding strategy, the collection of parameters X and the domain are converted to the bit string structure space S ;

Define the fitness function $f(X)$;

- (2) Determine the genetic strategy, including the choice of group size n , selection, hybridization, mutation, and determine the crossover probability p_{cr} , mutation rate p_m and genetic parameters;
- (3) Initialization the generate groups P randomly;
- (4) Calculate the group fitness $f(X)$ of the individual whose bit string were decoded;
- (6) In accordance with the genetic strategy, we use selection, hybridization and mutation operator to act on the groups to form the next generation of groups;
- (7) Judge whether the group performance is satisfied a target, or has completed the predetermined number of iterations, if not satisfied, return to step (6), else to modify the genetic strategy to return to step (6).

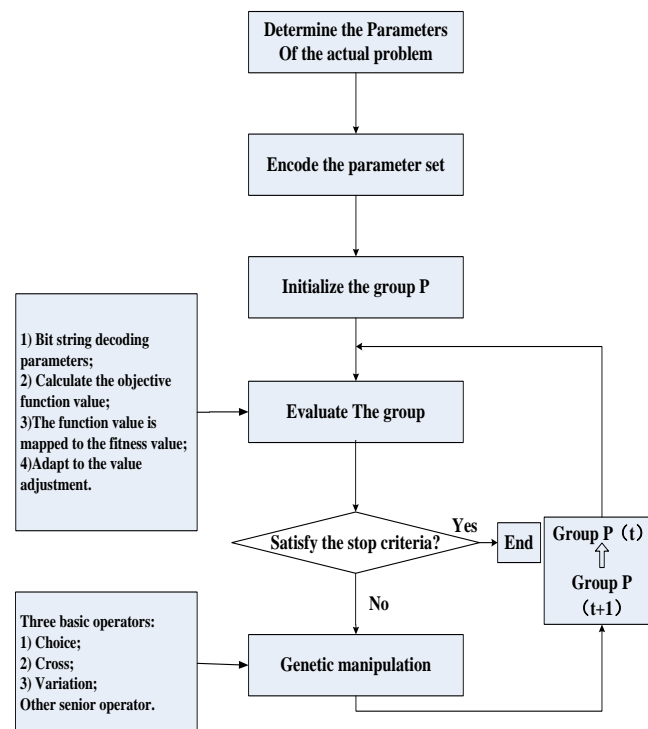


Figure 1. The basic flow diagram of the simple genetic algorithm

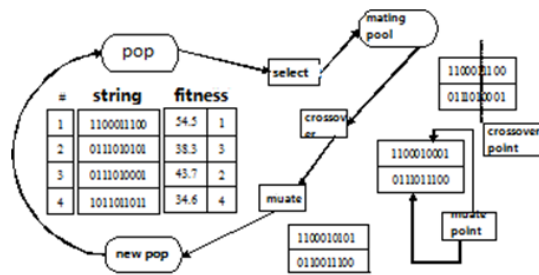


Figure 2. An simple example

3. Problem description and the example

This section focuses on how to use genetic algorithms to describe the urban residents' choice of travel mode.

Residents' choice of travel mode is facing a dynamic environment. If one is very familiar with the route, he/she can make quick decisions with the the real-time traffic, the travel time restrictions and cost constraints taken into consideration. However, as for the unknown route, the residents are not likely able to make the best decisions at the outset and can not find the optimal way to travel, we say the travel mode of decision-making process is a learning process of the residents. Residents choose a way to travel based on their past experience, as well as other residents travel circumstances, and then produce payment. The residents continue to accumulate experience, dynamically adjust their travel strategies step by step through imitation, trial and error optimization, so to make the next trip increase effectiveness. After a period of evolution, the adoption of a travel behavior will be whether or not caused by the increase or decrease the user profitable, so that make the user behavior strategy for distribution become in accordance with the principle of "survival of the fittest" evolution. Bounded rationality user policy learning and adjustment strategy equilibrium will lead the group on. On the micro level, individual residents constantly adjust their own way to travel, resulting in a more favorable payment in evolutionary biology, genes or culture can be copied, those who produce more payment will be imitated, but those who can not produce better payment would be eliminated (Hisashi, Christoph, Erez, et al., 2006). On the macro level, the proportion of group members adopt a specific strategy to become stable.

Studies have shown that travel time, cost and user income have great influence on travel mode choice (Yao et al., 2006).

This section gives an example of real life to describe how urban residents choose the travel mode based on genetic algorithm.

Scenario Description: A teacher from the Beijing Institute of Architectural Building (hereinafter referred to as Building) goes to the Golden Resources Shopping Mall (hereinafter referred to as the gold source) to have a class on Sunday afternoon. The class ends at 16:30 and the next begins at 18:00. Due to the returning of objects, and making ensure all of the students leave the building safely, the time to go after school is about 16:50. And this time is about the evening, if the teacher doesn't have dinner, the evening class is likely to be subject to impact, so it is best to set aside time for dinner, however, have dinner before the evening class is not necessary. Besides, the teacher must sign before 17:45. As the foregoing analysis, the teacher must make sure that the travel time is limited within one hour.

Current road conditions: If the teacher takes the bus, in accordance with the experience, it needs at least 50 minutes to reach the designated stations, and needs to walk for some time to get off to reach the destination, so there is certainly no time to eat dinner. What's more, that period is the sections of the rush hours period and the peak sections. But the biggest advantage to choose riding the bus is that costs are particularly low. If the teacher chooses to take a taxi, according to experience, he/she can reach the specified destination downstairs in at least 15 minutes, there is enough time to eat dinner, and a taxi can select the route freely, unlike the bus route is fixed and can not be changed, it means that the taxi will not be in the traffic jams. Choosing the taxi's biggest weakness is that the cost is 50 times over the bus. However, taking the ratio of delivery and taxi costs into account, the cost of a taxi is acceptable.

Firstly, convert the above scenarios into the model based on the genetic algorithm. The teacher's decision-making include:

- (1) Way to travel: bus or taxi?
- (2) Location: bus stop, or other?
- (3) Dinner: eat or not eat?

The purpose is to find a combination of these three decisions to produce the highest utility.

There are three decision variables, each variable can be assumed as one of the two possible values, so each possible decision-making of the practical problems can be represent with the characteristics of string of an alphabet whose length's $L = 3$ and size's $K = 2$ very natural. For each variable, the value 0 or 1 is specified as two possible options; in this search it includes $2^3 = 8$ possible decision alternatives. String length ($L = 3$), mapping of the alphabet scale ($K = 2$) form on the representation of the actual program. And the mapping decides the provisions of the specific location of the decision variables as the string of 0 or 1, the first step in the use of genetic algorithm for solving this problem is to select an appropriate representation scheme.

According to the above-described methods, Table 1 shows the four of them.

Program number	Travel mode	Start Location	Dinner	Binary representation
1	bus	Bus stop	Eat	11
2	bus	Bus stop	Not eat	10
3	bus	Else	Eat	1
4	taxi	Else	Not eat	100

Table 1. Representation of the travel decision-making

Assuming that this is the first time this teacher goes from the building destined to the gold source, and does not know the roads, the time and do not know how much maximum satisfaction utility the best decision gets and it can incur how much losses if made the wrong decision. Moreover, he/she does not even know which variable change separately will produce the effectiveness of the biggest changes.

Meanwhile, the teacher does not know whether he/she can get close to the global optimum, in the following process, we change a variable to select a good result every time, then a similar change in another variable, and pick the good result. In other words, the teacher does not know the variable whether can be a separate optimization, or whether they are interrelated and highly nonlinear.

We will make the four decision-making programs in Table 1 as the initial random population of the teachers' decision-making.

In fact, the teacher makes the selection of the decisions just the same way like the genetic algorithm. The start of the implementation of the genetic algorithm is to learn information about the environment through detecting the randomly selected points in the search space. In particular, from generation 0 (the initial random generation), the composition of the initial population is randomly generated in the individuals of the genetic algorithm. In this instance, the population size N is equal to 4.

In the genetic algorithm, the individuals in each generation group must detect in an unknown environment in order to get their fitness, then fitness value is taken as satisfied with the utility function value. In this instance, the initial group of four individual fitness value is given in Table 3 and the fitness is defined as follows.

Based on experience and personal preferences of the teacher, we define the travel satisfaction utility function of this teacher as the function of the travel time and the cost.

$$f(x_{ij}, y_{ij}) = \sum_{i=1}^3 x_{ij}^2 + \sum_{i=1}^3 y_{ij} + \varphi \quad (1)$$

$F(x_{ij}, y_{ij})$ is the sum of the travel time and the cost corresponds to the decision-making program of this teacher. The most important factor in this case is the travel time, so considering its position in the entire satisfaction of the effectiveness, we make it show as the form of a square. However the cost is not important than the time, we simply include it in the total utility directly. And φ represents as the extra satisfaction, it means that having the dinner has an extra influence on the teacher. What's more, φ is a constant, we define it equals to 50. Besides, i means the decision-making variables classification, 1 is for travel mode, 2 is for location and 3 is for dinner; and j represents 0-1 variables. x is the satisfaction utility corresponds to the travel time, and y is for the satisfaction utility corresponds to the cost.

Table 2 gives the satisfaction utility every decision gets corresponds to the travel time and cost (The value is from 1 to 10).

Decision variables	Travel time (x)	Cost (y)
Taxi (1)	10	9
Bus (0)	1	10
Bus stop (1)	3	2
Else (0)	4	2
Eat (1)	9	9
Not eat (0)	10	10

Table 2. The reference of the satisfaction utility

Among them, the travel time of choosing a bus is five times that of a taxi. And in this instance, choosing a bus will take a long time on the road, and it may be late for the class, so the taxi's satisfaction utility about the travel time is 10 points, the bus is 1 point. The changes of location have little effect on travel time and the cost, so there is little difference between the decisions. Although it is very good not to eat for travel time and the costs of satisfaction utility theory, but having the meal can make the night work smoothly. Reflections From a practical point of view, we should reduce the distance of the satisfaction utility for travel time and the cost.

We find the specific fitness value of the four special points (decision-making) by the detection of four random decision-making in the search space (utility value). In particular, we can see that the best individual 100 in the groups of generation 0 has produced the satisfaction of the utility value as 237, the worst individual satisfaction of the utility value of 010 is 132.

Generation 0		
k	String (x _k)	The fitness f(x _k)
1	11	162
2	10	132
3	1	169
4	100	237
sum		700
minimum		132
mean		175
maximum		237

Table 3. The fitness for the decision-making programs in the initial population

The only information that is actually used in the genetic is individual's fitness in the population. Through the simulation of biological natural selection and natural genetic process, the genetic algorithm transforms a group into a new group. A simple genetic algorithm consists of three genetic operators: replication, crossover and mutation.

The copy operator copied the current best individuals in the group to the new group according to the probability which was produced by their fitness divide the sum of the fitness. The sum of the fitness in generation 0 is 700. And because the best individual 100's value is 237, so the group's value which is attributed to the individual part 100 is 0.34. According to the fitness proportional selection, we expect the string 100 will appear in the new group for twice. Because the genetic algorithm is random, so the string 100 may occur 1 or 3 times in the new group, and even part of the tiny possibility of 4 times or not at all in the group is possible. At the same time, the individual 011, 010, 001 respectively attributes to the group as 0.23, 0.19, 0.24, so similar expectations of individual 011 and 001 in the emergence of new group respectively, and the individual 010 will disappear in the new group.

If the four strings copied to the new group is exactly accurate according to their expectations, string 100, 011, 001, 010 will appear twice, once, once and zero. The method to select the above series of copy process can be the choice of gamble which is commonly used, the basic steps are as follows:

- (1) Sum up the fitness of every string in the group;
- (2) Produce a random number between 0 and the sum of m;
- (3) From the string numbered 1 in the group, and sum its fitness with the subsequent string value together, until a sum equals to m or greater than m.

Table 4 shows an example of a gamble to choose, the first line are four numbers randomly selected from 0 to 700 (we use the Mat-lab to select these random numbers as follows: $700 \times \text{rand}$). For each random number, the number is cumulated begins with number 1 from Table 3 until the sum is equal to or greater than the random number generated by the beginning of the string, then we stop summing, and the last one into the string is the selected string.

Random number	123	655	642	474
The selected string	11	100	100	1

Table 4. The choice according to gamble

The result of gamble choice is to return a randomly selected string. Although the selection process is random, but each string of the opportunity to choose is referred to its adaptation proportional to the value. Of course, the worst string in the group sometimes may be selected because of the random of chose, which would affect the implementation of the results of the algorithm, but with the evolution process, this contingency will be negligible .

Table 5 gives a possible outcome of this new group generated by the reproduction operator with the chose of gamble. And this new group is called the mating pool. Mating pool in the middle between the current generation and the next, its size is N.

k	Generation 0				The mating pool
	String (x_k)	Fitness $f(x_k)$	$\frac{f(x_k)}{\sum f(x_k)}$	String	$f(x_k)$
1	11	162	0.23	11	162
2	10	132	0.19	100	237
3	1	169	0.24	100	237
4	100	237	0.34	1	169
Sum		700			805
Minimum		132			162
Mean		175			201
Maximum		237			237

Table 5. The mating pool after copy

The effect of the reproduction operator is to improve the average fitness value of the group. The average fitness value of the mating pool group is 201, and its starting point is only 175, and the worst individual adaptation of the mating pool is 162, while the worst individual in the initial population is 132. Low fitness individuals tend to be eliminated, while the high-fitness individuals tend to be copied, so representative in copy groups in the computation of these improvements, but at the expense of the loss of diversity of the population. The reproduction

operator does not produce a new individual, of course, the best individual's fitness will not improve in the group.

Genetic crossover operator can generate new individuals in order to detect a new point in the search space. The reproduction operator works on an individual each time, but the crossover works on two individuals in the mating pool randomly selected. The crossover operator generates two offspring strings, they are generally different from its parent string, and differ from each other, and each offspring the string contains the genetic material of two parent strings.

Table 6 shows the generation of a likely outcome of the application of copy operator and crossover operator from generation 0. The crossover probability is set to 50%, half of the mating pool's individuals (two) will be hybrid, the remaining individuals will only replicate.

K	Generation 0			The mating pool		Generation 1		
	String (x_k)	Fitness $f(x_k)$	$\frac{f(x_k)}{\sum f(x_k)}$	x_k	$f(x_k)$	Crossover point	x_k	$f(x_k)$
1	11	162	0.23	11	162	2	10	132
2	10	132	0.19	100	237	2	101	267
3	1	169	0.24	100	237	—	100	237
4	100	237	0.34	1	169	—	1	169
Sum		700			805			805
Minimum		132			162			132
Mean		175			201			201
Maximum		237			237			267

Table 6. The results of the copy and crossover operators

Hybridization of the two individuals happen to be 011 and 101, the crossover point is point two, and the two offspring strings are 010 and 101. According to Table 6, the last one is generation 0 in this genetic algorithm.

It can be seen from the table that the best individual of the first generation has a fitness value of 267, while the best individual in generation 0 only has the fitness value of 237. The hybrid produced some of the new individuals, and have higher fitness than its parent strings. Compared with generation 0 and generation 1, we can find:

- (1) The average fitness changed from 175 to 201;
- (2) The fitness value of the best individual changed from 237 to 267.

The above describes the genetic algorithm process from generation 0 to generation 1, then the iteration of genetic algorithm to perform this process, until it meets the stopping criteria-the largest evolution algebra.

4. Conclusion

In this instance, the final optimal decision-making program is the first generation of 101: taxi, bus stop and having supper. If we happen to know that 267 is the most highly satisfied with the utility, in this instance, we can stop in the first generation implementation of the genetic algorithm. When the genetic algorithm stops execution, put the current generation of the best individual to specify the results of the genetic algorithm. Of course, the genetic algorithm is generally not performed as in this case to generation to stop, but to dozens of generations, hundreds of generations or more generations.

Acknowledgments

This paper is supported by the National Natural Science Foundation (No. 71103014) and the Fundamental Research Funds for the Central Universities (No. 2011JBM032).

References

- Hisashi, O., Christoph, H., Erez, L., et al. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441, 502-505. <http://dx.doi.org/10.1038/nature04605>
- Ma, C., Wen, J., Li, C., & Li, X. (2009). Big city residents' travel mode decision-making method. *Journal of transportation engineering and information*, 2009(02).
- Song, D., Zhu, Y., & Wu, H. (2009). Interactive Genetic Algorithm based on the model of reasoning methods group. *Journal of Engineering Science*, 2009(11).
- Xu, Y., Li, X.H., Zhu, Y., & Shi, S. (2005). Satisfaction criteria of urban residents travel mode choice model. *Computer and Communications*, 2005(04).
- Yao, L., Guan, H., & Sun, L. (2006). Public transportation options to influence factors. *Dalian university press*, 239-243.

Journal of Industrial Engineering and Management, 2013 (www.jiem.org)



El artículo está con Reconocimiento-NoComercial 3.0 de Creative Commons. Puede copiarlo, distribuirlo y comunicarlo públicamente siempre que cite a su autor y a Intangible Capital. No lo utilice para fines comerciales. La licencia completa se puede consultar en <http://creativecommons.org/licenses/by-nc/3.0/es/>